

RESEARCH ARTICLE

Challenges in the selection of atmospheric circulation patterns for the wind energy sector

Verónica Torralba¹  | Nube Gonzalez-Reviriego¹ | Nicola Cortesi¹ |
Andrea Manrique-Suñén¹ | Llorenç Lledó¹ | Raúl Marcos¹ | Albert Soret¹ |
Francisco J. Doblas-Reyes^{1,2}

¹Barcelona Supercomputing Center (BSC),
Barcelona, Spain

²ICREA, Barcelona, Spain

Correspondence

Verónica Torralba, Barcelona
Supercomputing Center (BSC), Jordi
Girona, 29, 08034 Barcelona, Spain.
Email: veronica.torralba@bsc.es

Funding information

Horizon 2020 Framework Programme,
Grant/Award Number: GA 776787;
Ministerio de Ciencia e Innovación,
Grant/Award Numbers:
CGL2017-85791-R, IJCI-2016-29776

Abstract

Atmospheric circulation patterns that prevail for several consecutive days over a specific region can have consequences for the wind energy sector as they may lead to a reduction of the wind power generation, impacting market prices or repayments of investments. The main goal of this study is to develop a user-oriented classification of atmospheric circulation patterns in the Euro-Atlantic region that helps to mitigate the impact of the atmospheric variability on the wind industry at seasonal timescales. Particularly, the seasonal forecasts of these frequencies of occurrence can be also beneficial to reduce the risk of the climate variability in wind energy activities. *K*-means clustering has been applied on the sea level pressure from the ERA5 reanalysis to produce a classification with three, four, five and six clusters per season. The spatial similarity between the different ERA5 classifications has revealed that four clusters are a good option for all the seasons except for summer when the atmospheric circulation can be described with only three clusters. However, the use of these classifications to reconstruct wind speed and temperature, key climate variables for the wind energy sector, has shown that four clusters per season are a good choice. The skill of five seasonal forecast systems in simulating the year-to-year variations in the frequency of occurrence of the atmospheric patterns is more dependent on the inherent skill of the sea level pressure than on the number of clusters employed. This result suggests that more work is needed to improve the performance of the seasonal forecast systems in the Euro-Atlantic domain to extract skilful forecast information from the circulation classification. Finally, this analysis illustrates that from a user perspective it is essential to consider the application when selecting a classification and to take into account different forecast systems.

KEYWORDS

C3S seasonal forecasts, ERA5 reanalysis, Euro-Atlantic atmospheric patterns, *k*-means clustering, wind energy

1 | INTRODUCTION

Climate information and predictions have shown high potential for their application in different socioeconomic sectors such as energy (Garcia-Morales and Dubus, 2007; De Felice *et al.*, 2015; Reyers *et al.*, 2015; 9 Torralba *et al.*, 2017), agriculture (Ceglar *et al.*, 2017), transport (Palin *et al.*, 2016) or health (Lowe *et al.*, 2016). Nevertheless, the lack of tailored climate information that can be easily integrated in different decision-making processes has prevented a higher penetration of this information source. To address this gap, the development of robust methodologies that can be employed for the generation of products tailored to specific users' needs is required (Soret *et al.*, 2019). In this framework, climate sensitive sectors such as energy (Grams *et al.*, 2017; van der Wiel *et al.*, 2019) or water management (Lavaysse *et al.*, 2018) have shown interest in understanding the usefulness of atmospheric circulation patterns as an additional tool to the current climate information available to guide decision-making processes.

Wind energy users need to properly characterize the climate variability in a wide range of timescales (short-term, subseasonal, seasonal and decadal) because this variability can affect the balance between wind energy production and demand (Brayshaw *et al.*, 2011; Thornton *et al.*, 2017; Staffell and Pfenninger, 2018). The strong association between atmospheric circulation and wind power production suggests that the most recurrent large-scale atmospheric circulation structures can be used as a tool to understand the fluctuations of the wind energy generation (Curtis *et al.*, 2016; Zubiate *et al.*, 2017; Walz *et al.*, 2018). The occurrence of identifiable atmospheric patterns can be used to understand anomalies of the wind energy resources in the past months (forensic analyses) but also to make simplified estimations of the seasonal evolution of the atmospheric variability that could be useful to anticipate revenues and potential cash-flow problems, plan maintenance operations or anticipate supply-demand balance risks. Consequently, wind energy users have shown interest in knowing the influence of the atmospheric patterns on the wind energy resources as this can help to guide new deployment strategies that minimize the risk related to the variations in wind energy outputs (Grams *et al.*, 2017). Particularly, these users are interested in the impact of the atmospheric circulation on wind speed and temperature. These two variables are important because wind speed provides information about the available wind energy resources and temperature is directly linked to the energy demand (Bessec and Fouquau, 2008; De Cian *et al.*, 2013; De Felice *et al.*, 2015). These are relevant aspects for the wind energy community because one of the challenges

for renewable energy sources is to balance supply with demand.

Seasonal forecasts of the frequency of occurrence of specific circulation patterns would be of great value to wind energy users as this information can be used to understand changes in the essential climate variables (e.g., wind speed or temperature), which is important for the development of some strategies to mitigate the atmospheric variability related risks affecting wind energy resources. Therefore, wind energy users can combine seasonal forecasts of wind speed and temperature with forecasts of the expected frequency of occurrence of specific atmospheric patterns to take better informed decisions. This paper examines the capability of five Copernicus Climate Change Service (C3S) seasonal forecast systems to predict the average frequency of occurrence of these atmospheric patterns for the upcoming season. It is important to mention that the seasonal forecasts do not seek to predict the exact time of occurrence of the transitions between one pattern to another, as these transitions occur at a shorter timescale. Instead, these forecasts can be used to estimate which atmospheric patterns are the most likely to prevail over the coming season (Palmer and Anderson, 1994).

The most recurrent and persistent large-scale atmospheric circulation structures that allow the characterization of the complex atmospheric dynamics in the Euro-Atlantic region, which is a key region for several industrial activities, have already been studied in the scientific literature (Vautard, 1990; Michelangeli *et al.*, 1995). These patterns have been extensively used to investigate the atmospheric variability in the mid-latitudes, as they are associated with extreme weather events such as heat waves or droughts (Yiou *et al.*, 2008; Quesada *et al.*, 2012). The classification of the atmospheric circulation in a set of a few recurrent atmospheric patterns is a task that can be performed in different ways (see Huth *et al.*, 2008; Hannachi *et al.*, 2017, for comprehensive reviews on the topic). In this work, an unsupervised classification of the daily circulation into clusters with similar spatial patterns has been performed by the *k*-means algorithm (Hartigan and Wong, 1979), which is one of the most common machine learning approaches applied in climate research (Michelangeli *et al.*, 1995).

The wind industry has shown interest in the relationship between the atmospheric patterns and the fluctuations in energy production and demand. However, the potential of these atmospheric patterns for the development of a climate service has not been widely exploited yet. One of the reasons is that most of the existing studies focus mainly on the winter season (e.g., Cassou, 2008; Dawson *et al.*, 2012; Stryhal and Huth, 2017) and these classifications for one specific season are not sufficiently

detailed to fully understand the variability of user-relevant variables throughout the year (Grams *et al.*, 2017). Recently, some studies have attempted to overcome this limitation by proposing yearly-round classifications that are defined for each month or the full year (Vrac *et al.*, 2014; Grams *et al.*, 2017; Cortesi *et al.*, 2019) but the ability of the seasonal forecasts to simulate the variability of these patterns remains largely unexplored.

The main goal of this paper is to identify those atmospheric circulation patterns that can be used for the development of a climate services' product for the wind energy sector to complement the seasonal forecasts of climate variables that are directly produced by the seasonal forecast systems. As there is no standard procedure available in the literature to define these atmospheric patterns for the four climatological seasons and, at the same time, taking into account these specific needs, the criteria needed to generate a product that can satisfy the wind energy requirements have been assessed. One of the major decisions involved in the definition of these user-oriented atmospheric patterns through the *k*-means analysis is the choice of the optimal number of clusters. Although it has been shown that four clusters are a good choice for the Euro-Atlantic region in winter (Michelangeli *et al.*, 1995; Ferranti and Corti, 2011), the optimal number of clusters for the other climatological seasons has not been discussed so far. The four clusters obtained from the *k*-means algorithm in winter are usually referred to as weather regimes. These four atmospheric patterns are the positive and negative North Atlantic Oscillation phases, the atmospheric blocking and the Atlantic ridge (Cassou, 2008). In this work, the cluster analysis has been applied for a number of patterns equal to three, four, five and six clusters per season. This is different from the four clusters traditionally obtained, for this reason, the term "weather regimes" has not been used in this manuscript. The benefits and drawbacks of these classifications for each specific number of clusters have been discussed taking into account three aspects that are relevant for the development of this climate services' product: (a) the spatial robustness of the patterns (b) the usefulness of each classification to understand the variability of two essential climate variables for the wind industry (wind speed and temperature) and (c) the skill of the seasonal forecast systems in predicting the frequency of occurrence of these clusters. The goal is to retain a reduced number of patterns that can be easily identifiable and understandable by the users that reproduces the year-to-year fluctuations of the wind speed and temperature and that can be predicted by the seasonal forecast systems.

This paper is organized as follows. The details on the data processing and methodology have been included in Section 2. Section 3 presents the main results and discusses their relevance in the user context. The main

conclusions of this work and their relevance to provide a climate services' product are explained in Section 4.

2 | DATA AND METHODOLOGY

2.1 | Datasets

The ERA5 reanalysis (C3S, 2017) has been used as an observational reference for the identification of a reduced set of atmospheric patterns with a cluster analysis. ERA5 is the latest ECMWF atmospheric reanalysis, and it is available for the period between January 1979 to the present. It is based on a version of the ECMWF atmospheric model that was operational in 2016 and employs a four-dimensional variational analysis (4D-Var) for data assimilation. This reanalysis has a spatial resolution of approximately 31 km. The sensitivity of the classifications to the choice of the reanalysis has been already explored (Stryhal and Huth, 2017; Cortesi *et al.*, 2019) and it has been demonstrated that different reanalyses produce equivalent classifications in most of the months. The ERA5 data have been bilinearly interpolated to a 1° regular grid. The atmospheric patterns computed in this work are based on daily means of sea level pressure (SLP) in the period 1979–2016 (38 years) for the Euro-Atlantic region [$27^\circ - 81^\circ$ N, 85.5° W– 45° E]. This domain has been also employed in Cortesi *et al.* (2019) and it is similar to other domains already employed in the literature (e.g., Dawson and Palmer, 2015; Ferranti *et al.*, 2015). Nevertheless, the choice of a specific domain for the definition of the clusters could have an impact on the results, particularly when these clusters are used to understand specific climate variables. For example, Beck *et al.* (2016) showed the differences in the optimal domain between winter and summer seasons and also when these clusters are used to understand the impact of the large-scale circulation on temperature or precipitation. For that reason, the development of a specific service for wind energy users interested in specific locations could imply the modification of the domain.

Although geopotential height at 500 hPa has been widely used in the literature to obtain a set of clusters in the Euro-Atlantic region (Cassou, 2008; Dawson *et al.*, 2012; Ferranti *et al.*, 2015) this assessment is based on SLP (Fereday *et al.*, 2008; Neal *et al.*, 2016; Stryhal and Huth, 2017). This choice is justified because SLP is less affected by global warming than the geopotential height (e.g., Hartmann *et al.*, 2013) and can provide more information about the impact of the circulation on surface variables relevant for wind energy applications. To understand the influence of the atmospheric circulation on user-relevant variables, the daily means of the 10 m

wind speed and 2 m temperature from the ERA5 reanalysis have been reconstructed using the circulation patterns. We have used the methodology described in Cortesi *et al.* (2019) and detailed in Section 2.3. Daily anomalies have been computed as deviations from the daily climatologies for each individual grid point. The daily climatologies have been previously smoothed out by a Loess filter (Cleveland and Devlin, 1988) to remove the short-term variability (Mahlstein *et al.*, 2015). The SLP anomalies have been weighted by the cosine of the latitude before performing the cluster analysis to take into account the different areas of each grid box.

Seasonal forecasts from the five C3S different forecast systems listed in Table 1 have been employed in this work. These forecasts span 6 months into the future and they are produced by different centres: ECMWF (European Centre for Medium-Range Weather Forecasts), Met Office, Météo-France, CMCC (Centro Euro-Mediterraneo sui Cambiamenti Climatici) and DWD (Deutscher Wetterdienst). These centres have employed two different strategies for the generation of the ensemble members, the burst mode and the lagged method. The ECMWF, DWD and CMCC are producing ensembles in burst mode, which means that all the members are initialised simultaneously at the same time (they all share the start date), but from slightly different (perturbed) initial conditions. The Met Office and Météo-France have produced the ensemble members in a lagged mode, which means that the members are initialised on different dates during the month. The seasonal forecasts from the lagged ensembles have been reorganized taking as the first day the value that verifies on the selected common start date, which in the C3S case is the first day of the month, discarding any data before this date. For example, the ensemble with 25 members corresponding to the 1st of November in the Météo-France System 6 system (MF-S6) is the combination of one member initialised the first of November plus 12 members initialised the 25th of October, but only considered from the 1st of November, plus 12 members initialised the 20th of October, also considered only from the first of

November. Similarly, the ensemble from the Met Office GloSea5 for the same start date (i.e., the 1st of November) with 28 ensemble members corresponds to the combination of seven members initialised on four different start dates: 1st of November, 25th of October, 17th of October and 9th of October.

We have employed the SLP seasonal forecasts in the hindcast period 1993–2016 (24 years), which is the common period for all five systems as provided by C3S. This analysis focuses on the one-month lead forecasts for the four climatological seasons: DJF (December–January–February), MAM (March–April–May), JJA (June–July–August) and SON (September–October–November). This corresponds to predictions initialised in November, February, May and August, respectively.

2.2 | Classification methods

The *k*-means algorithm has been used to identify the clusters in ERA5. For the seasonal forecasts the assignment method based on the minimum root-mean-square distance was used to classify the SLP daily anomalies.

2.2.1 | K-means algorithm

One of the most common methods used in climate research for the classification of the atmospheric circulation is the *k*-means algorithm (Hartigan and Wong, 1979). This clustering method produces a partition of all daily SLP anomalies (data points) in a predefined number of clusters (*k*), which minimizes the within-cluster-sum-of-squares (i.e., the squared deviations from the data points to the cluster centroid) while maximizes the inter-cluster variance. The cluster partitions are obtained in an iterative process that can lead to slightly different classifications when the cluster analysis is repeated. Although the centroid coordinates are usually obtained by projecting the anomaly field onto empirical

Institution	System	Acronym ^a	Ensemble size	Reference
ECMWF	SEAS5	ECMWF-S5	25	Johnson <i>et al.</i> (2019)
Met Office	GloSea5-GC2	UKMO-GS5	28 ^b	Williams <i>et al.</i> (2015)
Météo-France	System 6	MF-S6	25 ^b	Dorel <i>et al.</i> (2017)
CMCC	SPSv3	CMCC-S3	40	Sanna <i>et al.</i> (2017)
DWD	GCFS2.0	DWD-S2	30	DWD (2020)

^aThis is the name used in the manuscript to refer to these systems.

^bThese ensemble members have been generated in a lagged mode instead of the burst mode employed by the other systems.

TABLE 1 Definition of the seasonal forecast systems employed in this work. These forecasts are available through the C3S Climate Data Store (CDS, <https://cds.climate.copernicus.eu/#!/home>)

orthogonal functions, here the k -means method has been directly applied over the daily SLP anomalies to take into account extreme SLP values (Cortesi *et al.*, 2019).

One of the major assumptions of this method is that the number of clusters (k) needs to be determined a priori. There are several methodologies with different complexity to identify the optimal number of clusters for the clustering of a specific dataset (Michelangeli *et al.*, 1995; Straus *et al.*, 2007). However, these methodologies have been mostly applied for the winter season in the Euro-Atlantic domain, where four clusters have been shown to be the best choice. Beyond winter, some difficulties have been identified. For example, in the summer season, it has been already shown that the atmospheric circulation over the Euro-Atlantic region shows less variability, and consequently an agreement in the classification is more difficult (Yiou *et al.*, 2008). One of the most basic approaches for the selection of the number of clusters is to run the k -means clustering for the different number of clusters several times. Following the Elbow criterion, the number of partitions should be chosen so that another cluster does not reduce the variance substantially (e.g., Jolliffe and Philipp, 2010; Gueye *et al.*, 2012). In this work, the k -means clustering has been applied with 30 initial centroids and 100 iterations for the ERA5 SLP data. The calculations were performed separately for each season. The variance in the k -means cluster analysis has been estimated as the ratio between the sum of the within-cluster-sum-of-squares and the total variance of the sample (daily anomalies). The results are illustrated in Figure 1 for the different seasons.

The variance for the different seasons decreases very slowly with the number of partitions. However, there are variations depending on the season. The slowest decrease of the variance is identified in JJA (Figure 1, red line) where it varies from 1 for $k = 1$ to 0.75 for $k = 7$. The results for SON are analogous to those identified in JJA, although the variance is slightly superior to that season for $k = 2, 3$, and 4. In the case of the boreal winter (DJF) the variance change from 1 to 0.75 is obtained from $k = 1$ to $k = 4$. This figure suggests that $k = 4$ might be a good option for winter and even for the boreal spring (MAM), but the optimal number for each specific season is not clear enough. Hence, due to an all-year-round classification being needed for the development of an operational climate service for the wind energy users, the implications of the selection of different numbers of clusters has been evaluated from different perspectives. In this paper, we do not attempt to discuss if the clusters obtained in the different seasons correspond to dynamical characteristics of the atmospheric circulation (Stephenson *et al.*, 2004; Christiansen, 2007). The main goal is the identification of a set of atmospheric patterns that can be

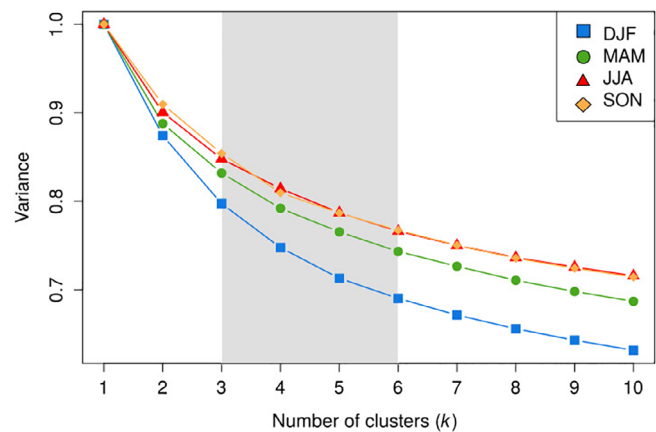


FIGURE 1 Variance as a function of the number of clusters (k) per season: DJF (blue), MAM (green), JJA (red), and SON (yellow). The variance is the ratio between the sum of the within-cluster-sum-of-squares and the total variance of the sample (daily anomalies). The k -means has been applied for the ERA5 SLP in the 1979–2016 period with 100 iterations and 30 initial centroids. The grey region highlights the values of k considered in this work

used to provide useful information to users. For that aim, we have considered four different numbers of clusters ($k = 3, 4, 5, 6$) to illustrate the challenges associated with this goal. A higher value of k might be considered, but that classification will contain very similar spatial structures, which would be very difficult to distinguish and predict by the seasonal forecast systems.

2.2.2 | Classification of the seasonal forecasts

Once the k -means method has been applied to define a set of clusters for the ERA5 reanalysis in each specific season, these clusters have been used for the definition of the atmospheric patterns in the seasonal forecasts. The assignment of each seasonal forecast of SLP daily anomaly to one of the defined clusters (separately for the classification with 3, 4, 5 and 6 partitions) from the ERA5 reanalysis has been performed by the minimum root-mean-square distance (RMSD) between the forecast anomalies and the cluster centroids (Neal *et al.*, 2016; Lavaysse *et al.*, 2018; van der Wiel *et al.*, 2019). The predicted patterns are obtained through the average of the SLP anomalies in all the days assigned to each specific pattern. The main advantage of this approach is that it guarantees resemblance between the atmospheric patterns obtained from the seasonal forecasts to those defined in the observational reference. The direct application of the k -means algorithm to identify a set of clusters

in the seasonal predictions considered in a previous work (Torralba, 2019) produces patterns that cannot be automatically mapped into those from the observational reference. The fact that the classification may not be the same in the forecasts and the observational reference makes the forecast quality assessment of the predicted atmospheric patterns impossible. As the forecast quality assessment is a crucial step in the development of any climate information, the clusters obtained from the application of the k -means methodology to the seasonal forecasts are not the most suitable product to be integrated into an operational climate service. Consequently, the use of the RMSD method ensures the matching between the forecasts and the observational reference, allowing the verification of the frequency of occurrence of each predicted pattern, and thus, the provision of a well-documented product to the users.

2.3 | Methods

Different products and metrics have been used to evaluate the clusters in terms of spatial patterns and frequency of occurrence.

The dissimilarity between the atmospheric patterns obtained for each k value has been measured in terms of spatial correlation (Jolliffe and Stephenson, 2012). It allows quantifying the degree of similarity among the spatial patterns. If two clusters show a high spatial correlation between them it is because these patterns are representing similar atmospheric conditions. In this case, a classification with a reduced number of patterns where these two patterns appear merged into one would be better. A one-sided t test has been applied to evaluate if the spatial correlation values are significant at the 95% confidence level. This assessment of the spatial patterns has been also performed in terms of the standard deviation ratio between the maps. This metric has been considered to take into account the similarity between patterns in terms of amplitude, as two maps with very high spatial correlations can show strong differences in their intensity.

To assess the effectiveness of the full set of clusters in describing the year-to-year variability of these climate variables a reconstruction methodology has been applied. Reconstruction methods have been already employed to quantify the influence of atmospheric circulation patterns in the wind speed and wind farms capacity factors (Cortesi *et al.*, 2019; Garrido-Perez *et al.*, 2020, respectively) and also in temperature and precipitation (Hall and Hanna, 2018). Furthermore, reconstruction methodologies have been already applied to investigate the role of the atmospheric

circulation on historical temperatures (Deser *et al.*, 2016) or the long-term evolution of the European precipitation (Fereday *et al.*, 2018).

We employ the reconstruction method to assess the relationship between the atmospheric circulation and the surface wind speed and temperature, estimating also how the different number of clusters affects the relationship. The reconstruction is based on the composite maps of wind speed and temperature and the frequency of occurrence (i.e., the percentage of days in a season and year) of each specific cluster. The reconstructed climate variables (varRecon) are obtained by the following equation:

$$varRecon_{sea,yr}(lat,lon) = \sum_{r=1}^R CM_{r,sea}(lat,lon) \cdot freq_{r,sea,yr} \quad (1)$$

where $freq_{r,sea,yr}$ is the percentage of days in a season and year assigned to a specific cluster and $CM_{r,sea}(lat,lon)$ is the composite map of the variable to be reconstructed for each pattern in a particular season. The reconstruction has been applied in leave-one-out cross-validation, which means that the anomalies corresponding to the year to be reconstructed are excluded from the computation of the composite maps.

To evaluate the performance of the reconstruction method, the Pearson correlation between the reconstructed climate variables and the climate variables obtained directly from ERA5 has been computed. Furthermore, Pearson correlation has been also employed to investigate the potential skill of the seasonal forecasts to simulate the frequency of occurrence of the atmospheric patterns. The statistical significance of the Pearson correlation is based on a t test that takes into account the autocorrelation. Note that the autocorrelation can lead to slightly different significance levels for equal correlation values.

3 | RESULTS

3.1 | Spatial consistency of the clusters

The dissimilarity between the clusters obtained for each number of partitions ($k = 3, 4, 5$ and 6) has been explored by a criterion based on the spatial correlation coefficient between the clusters not exceeding a certain threshold (Neal *et al.*, 2016; Grams *et al.*, 2017). Spatial correlations are shown in Figure 2 for each season. To complement the information of the spatial correlations by taking into account the differences in terms of amplitude among the patterns, the standard deviation ratio between the clusters have been computed (Figure S1).

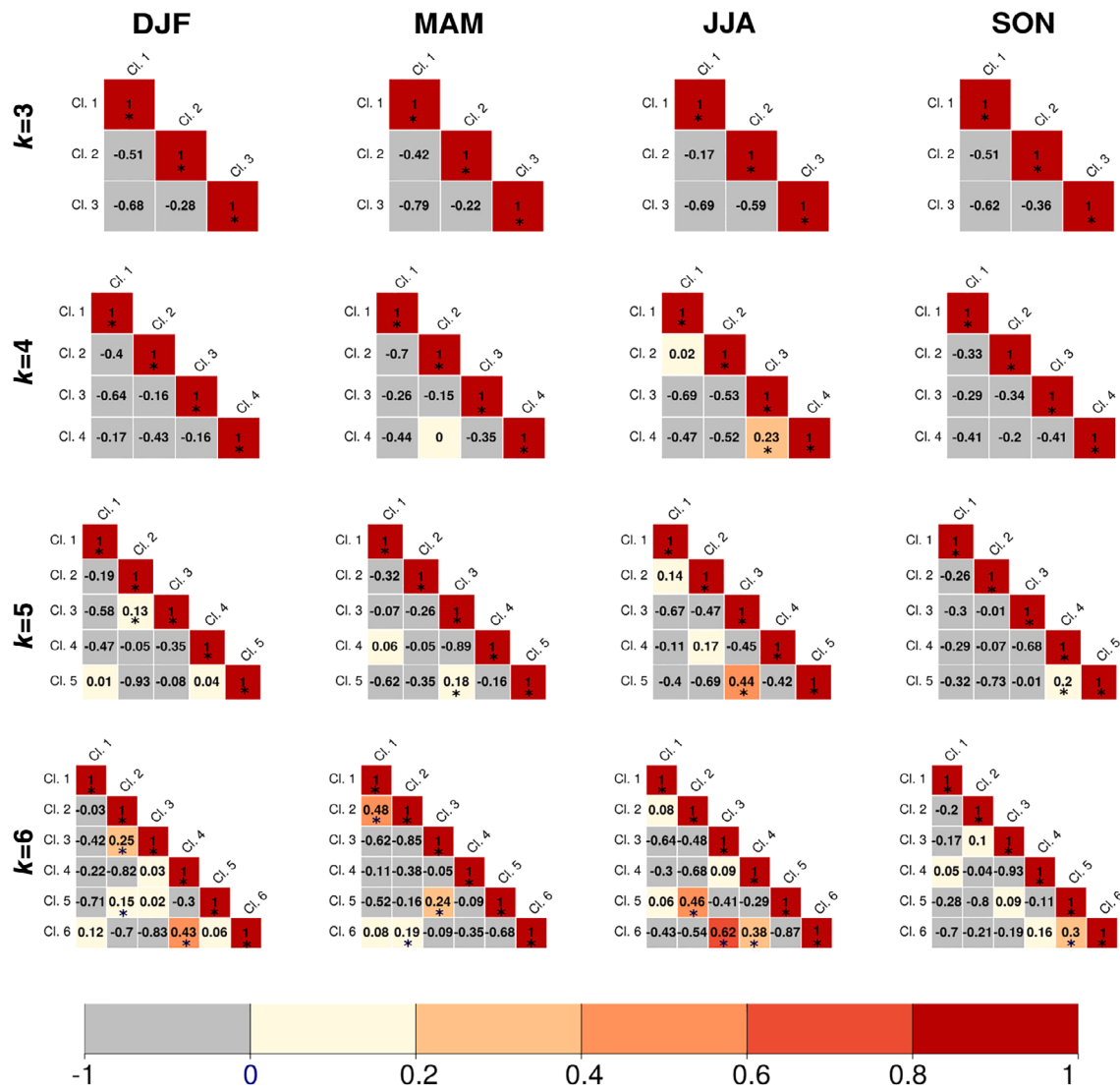


FIGURE 2 Spatial correlation between the clusters obtained by the application of the k -means method with $k = 3, 4, 5, 6$ and for the different seasons. The clusters have been computed for the daily anomalies of the ERA5 SLP in the 1979–2016 period. Asterisks denote significant spatial correlations at the 95% significant level (one-sided t test)

The patterns resulting from the application of the clustering method with $k = 3$ (Figure 2, first row) do not show positive spatial correlations among different clusters for any season, indicating that the three clusters do not have spatial correspondence. When we increase k from three to four (Figure 2, second row) analogous results are found for all the seasons, with the only exception of JJA, for which a significant correlation value equal to 0.23 indicates a certain degree of similarity between cluster 3 and cluster 4. These cluster 3 and 4 are also similar in terms of amplitude as the standard deviation ratio is 0.9 (Figure S1). This suggests that $k = 3$ might be a good choice for JJA.

The results of the k -means for $k = 5$ show positive and significant spatial correlations for different clusters in the four seasons. However, only in JJA, the spatial

correlation between clusters 3 and cluster 5 exceed 0.4 indicating a high resemblance between these two patterns. This resemblance is also seen in terms of amplitude, as the standard deviation ratio of these two maps is 0.99 (Figure S1). For DJF, MAM and SON, the positive correlations for different clusters are lower than 0.2, and the standard deviation ratio for those patterns show some deviations from 1, which might suggest that these five patterns could be considered in specific applications.

For $k = 6$, significant spatial correlations are obtained for the four seasons, indicating that at least one of the clusters is very similar to another cluster. For example, in DJF clusters 4 and 6, or in MAM clusters 1 and 2 are very similar with spatial correlations of 0.43 and 0.48, respectively. This reveals that six clusters might not be needed

to represent the atmospheric circulation over the Euro-Atlantic region in all the seasons.

To better illustrate this result, Figure S2 shows the DJF spatial patterns for $k = 3, 4, 5$ and 6. The composite maps for $k = 3$ in winter (Figure S2, first row) resemble the patterns of the positive phase of the North Atlantic Oscillation (NAO), the atmospheric blocking and the negative phase of the NAO. When the number of clusters is increased to four, a new pattern that resembles the Atlantic ridge can be identified. These four clusters have been widely discussed in the literature (Vautard, 1990; Cassou *et al.*, 2005). For $k = 5$, clusters 2 and 3 show a certain degree of similarity as shown by the significant spatial correlations obtained for this season, although this

correlation is only 0.13 (Figure 2). For $k = 6$, clusters 4 and 6 display comparable patterns, which suggests that the two patterns might not be necessary to describe the Euro-Atlantic atmospheric circulation in DJF. The spatial patterns of the clusters corresponding to MAM, JJA and SON can also be found in the supporting information (- Figures S3–S5, respectively).

3.2 | Reconstruction of user-relevant climate variables

One of the goals of this work is to classify SLP daily anomalies into clusters that can be used to characterize

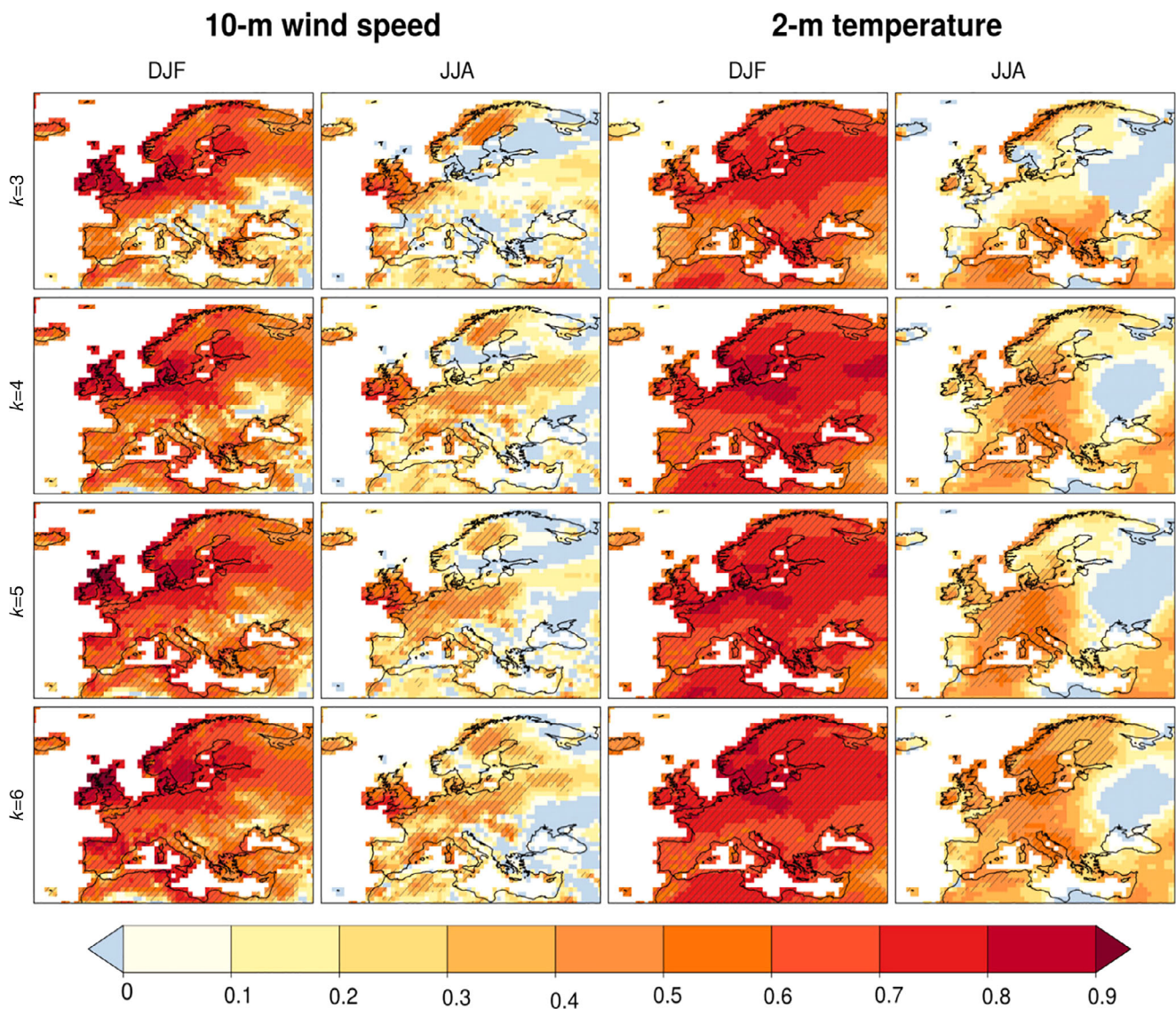


FIGURE 3 Pearson correlation between the reconstructed 10 m wind speed and 2 m temperature with the corresponding original variable obtained directly from ERA5 in the 1979–2016 period. These results have been obtained for the different seasons (columns) and for different number of clusters (rows). Hatched areas show regions where the correlations are significantly different from zero (one-sided t test at a 95% confidence level)

the impact of the atmospheric circulation on surface variables relevant for energy users. For that reason, it is important to take into account the ability of the classification to explain wind speed and temperature variations. In this context, the frequency of occurrence of each cluster has been used to reconstruct the mean wind speed and temperature for each specific season and year. This reconstruction is based on the methodology explained in Section 2.3. The reconstructed wind speed and temperature have been compared with the corresponding original variables obtained directly from ERA5 in terms of Pearson correlation of the DJF and JJA seasonal averages (Figure 3). The corresponding results for MAM and SON have been included in the supporting information (Figure S6). The figure only shows results for land areas and coastal zones because these are the regions of interest to the users, as they are locations where the wind farms can currently be installed.

High correlations are found between the reconstructed and the ERA5 wind speed for the different number of clusters considered in some regions in northern Europe, such as the British Isles, northern France, Germany, Denmark and also Scandinavia (Figure 3). This reveals that in adjacent regions to the North Sea, the atmospheric circulation represented by these clusters plays a dominant role in defining the wind speed inter-annual variability. However, it can be noted that the strength of this correlation shows an annual cycle with the lowest correlations obtained for the boreal summer (JJA, Figure 3, second column). For example, in southern Scandinavia, the JJA correlations are below 0.1. This reflects that either the atmospheric circulation is not dominating the wind speed variability in that specific season or that the classifications obtained are not representative of that region.

Figure 3 not only illustrates the differences between summer and winter clusters to efficiently reconstruct the wind speed, but also the sensitivity of this reconstruction to the number of clusters considered. The low correlations in JJA for $k = 3$ become significant in central Europe, and also in some regions in southern France or Italy when four clusters ($k = 4$) are employed, indicating that the four clusters better describe the variability of wind speed in that region compared to $k = 3$. The spatial correlation between cluster 3 and cluster 4 (shown in Figure 2), suggested that four clusters might not be the optimal number of clusters for JJA, being $k = 3$ a better option. However, four clusters show some added value when these clusters are used to understand the variability of the wind speed over Europe. This illustrates the importance of considering criteria beyond statistical metrics when designing a system to describe the atmospheric circulation variability in a climate services context.

The changes in the correlation obtained for the reconstructed wind speed for the different number of clusters have been illustrated in Figures S7 and S8. They show that the highest improvement in the reconstruction of wind speed is obtained when the number of clusters increases from three to four. Some exceptions can be identified for example in MAM when the correlations of the reconstructed wind speed based on five clusters improve in Scandinavia compared to the reconstruction based on only four clusters. However, as it has been shown in Figure 2, a classification based on six clusters would not be the optimal option because two clusters have similar patterns.

In the case of the reconstructed temperature (Figure 3, third and four columns) most of the European countries show positive and significant correlations exceeding 0.5 in the winter (DJF). This reveals that changes in the frequency of occurrence of the atmospheric clusters are linearly related to the temperature variability. However, in JJA, the reconstructed temperature in western Europe shows negative correlations with the corresponding variable from ERA5. This result suggests that there are additional factors to the North Atlantic large-scale atmospheric circulation affecting the variability of the temperature in this region.

The correlations of the reconstructed temperature in DJF show little variations among the classifications with 3, 4, 5 and 6 clusters. This can be also observed in the correlation differences between the temperature reconstructed with the classifications obtained for different k (Figure S7). However, the increase from three up to four clusters (Figure S7, row 1, column 3) shows a low but significant increase of the correlation indicating a general improvement when $k = 4$. In JJA, the correlation differences (Figure S7) show the highest values between $k = 3$ and $k = 4$, which show that the use of four clusters for the reconstruction of temperature is more efficient than three clusters. For MAM and SON (Figure S7), the highest improvement in terms of correlation is obtained when the number of clusters is increased from three to four, suggesting that a classification with four clusters could be the optimal option when these clusters are used to understand the temperature variability. However, a correlation increase is also identified for the reconstructed temperature over the British Isles when the number of clusters is increased from five to six in SON (Figure S8).

The reconstruction of wind speed and temperature from the frequency of occurrence of the classifications with different number of clusters suggests that there are several regions in Europe where atmospheric variability plays an important role in the user-relevant variables. Nevertheless, this approach is not equally effective for all

seasons. For example in boreal summer, SLP clusters are less efficient explaining the surface variability, as illustrated by the reduction of the correlation values between the reconstructed and the original variables.

The reconstruction method allows us to explore in which regions the clusters are most useful to understand the fluctuations of the wind speed and temperature. The classification obtained for four clusters has shown its ability to provide reconstructed variables with high correlations for the four seasons. However, this selection could be modified depending on the users' needs (e.g., depending on the location of their wind farms). For the development of a specific product for a user interested in a specific location, the cluster region could be also optimized to enhance the representation of the year-to-year fluctuations of the wind speed and temperature in each season. The same methodology has been applied to reconstruct wind speed at 100 m, as this level is close to the wind turbine heights than the 10 m considered above and can be also interesting for the wind energy users.

Figure S9 shows that reconstructed 100 m wind speeds have very similar correlations to those at 10 m, with only a few regions in which the wind speed at 100 m shows a stronger correlation with the atmospheric circulation patterns (e.g., southern Europe in MAM or central Europe in JJA).

3.3 | Skill in the seasonal forecasts of the frequencies

The first step in the development of any seasonal forecast product tailored to specific applications is the identification of the seasonal forecast system that can produce the best quality climate information. This is an important step, as the quality of these systems is highly dependent on the selected season and the variable. As a starting point, the ability of the C3S seasonal forecast systems in producing skilful SLP forecasts has been assessed with the Pearson correlation (Figure 4). The corresponding

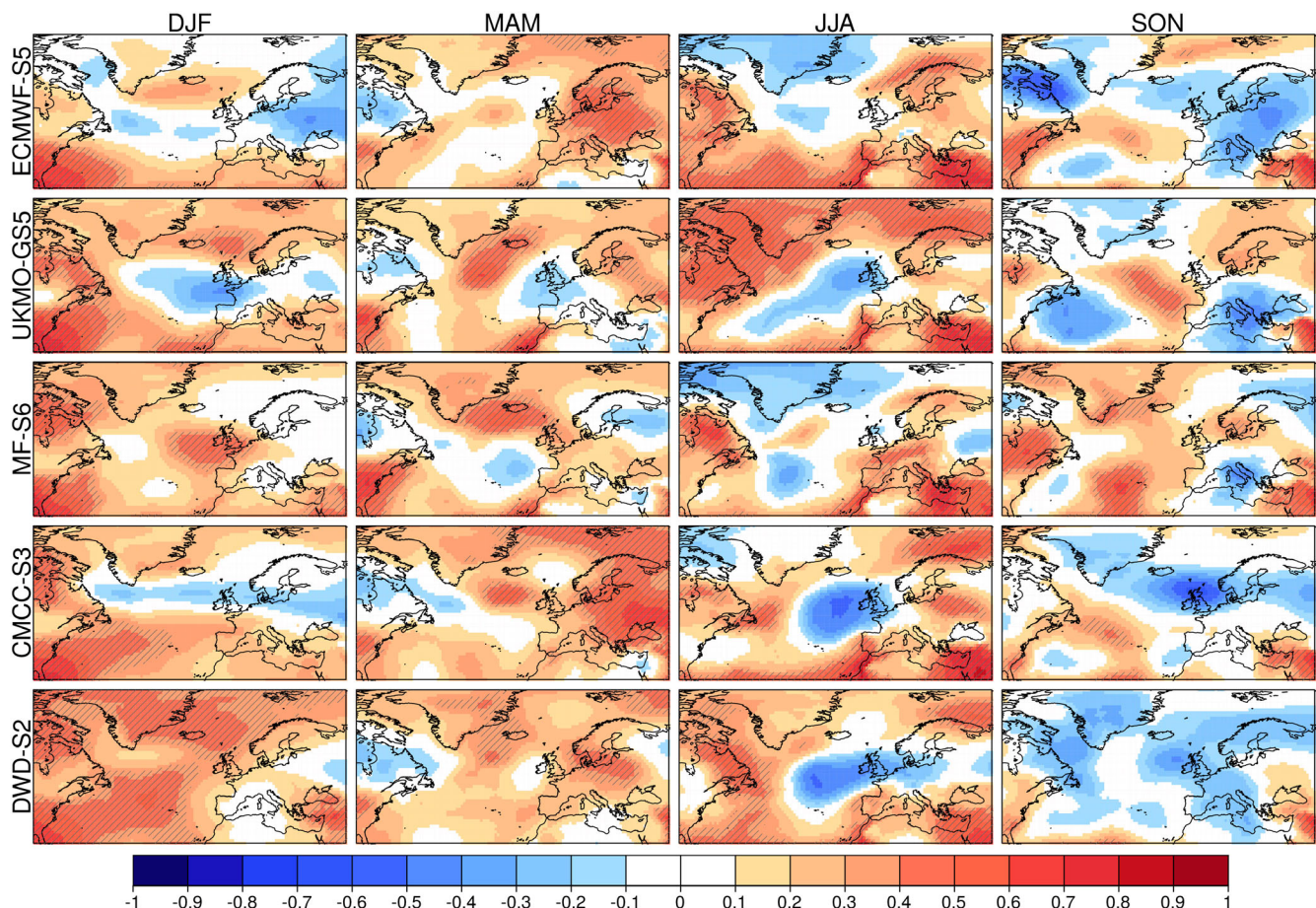


FIGURE 4 Pearson correlation of the SLP in the ensemble mean predictions from the different C3S systems (Table 1) and ERA5 in the 1993–2016 period. These results have been obtained for DJF, MAM, JJA and SON seasons (columns) corresponding to the seasonal predictions initialised one-month-ahead (i.e., November, February, May and August). Hatched areas denote regions where the correlations are significantly different from zero (one-sided t test at the 95% confidence level)

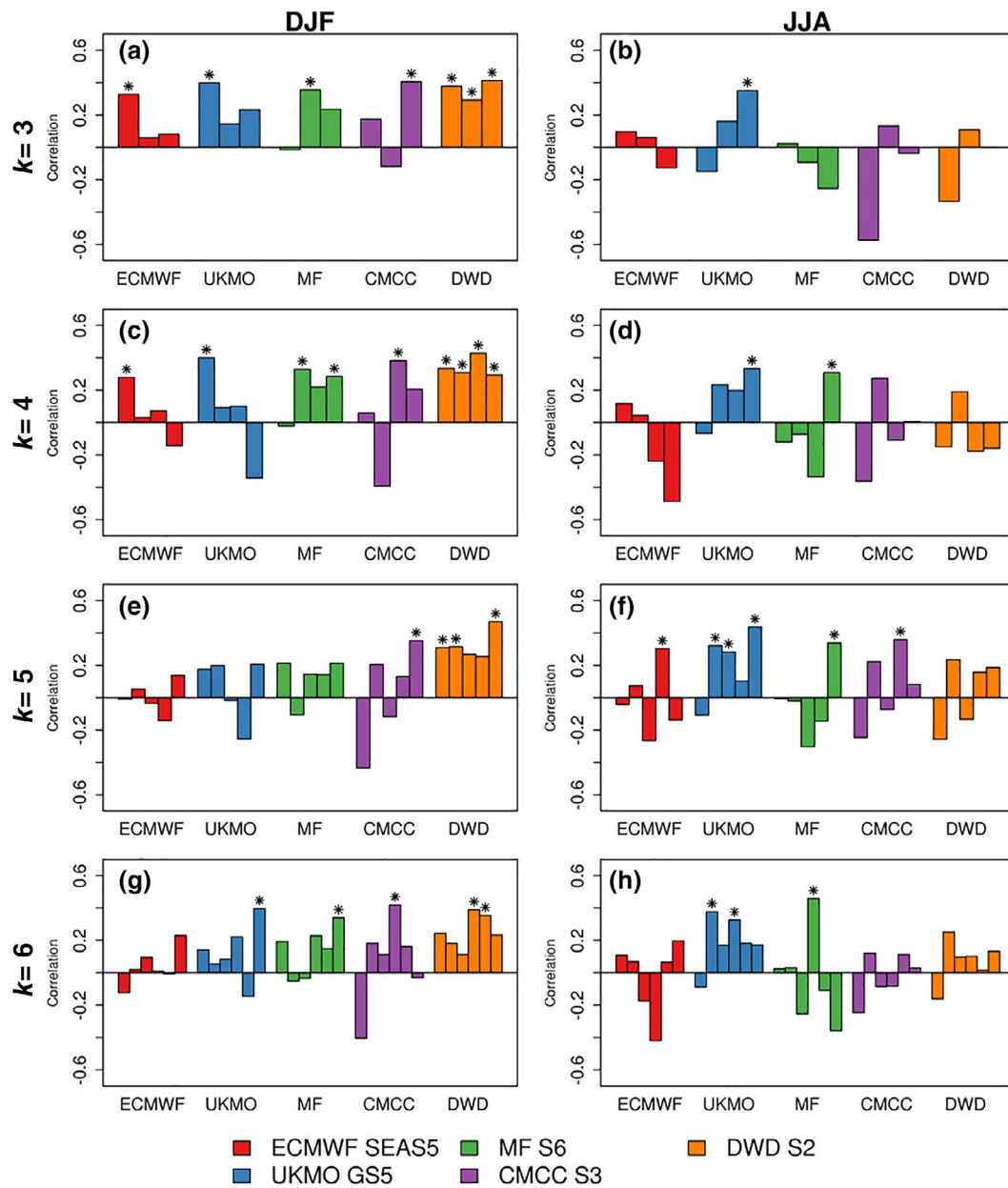


FIGURE 5 Pearson correlation between the C3S seasonal forecast of the frequency of occurrence and the corresponding frequencies for the ERA5 clusters (used as a reference). These results have been obtained for the different number of k (rows) and for DJF (panels a,c,e,g) and JJA (panels b,d,f,h) corresponding to the seasonal predictions initialised one-month-ahead (i.e., November and May) in the 1993–2016 period. Asterisks at the top of some bars indicate correlations that are significantly different from zero (one-sided t test at the 95% confidence level)

figures but for wind speed and temperature have been also included in the supplementary material (Figures S10 and S11), as this contextual information could be also relevant for the wind energy users.

Figure 4 illustrates the differences and the commonalities among the different C3S seasonal forecast systems in the correlation of the SLP with ERA5 over the Euro-Atlantic domain for the four seasons. In DJF (Figure 4, first column), the only seasonal forecast system providing positive correlations over most of the North Atlantic

region and eastern North America is the DWD-S2 system. Nevertheless, the correlation over the European continent is not significant. In MAM (Figure 4, second column), the five forecast systems agree on the positive correlations obtained for their SLP over southeastern North America and eastern Europe being the latest significant for a large area in both the ECMWF-S5 and CMCC-S3 systems. In JJA (Figure 4, third column) the five systems considered show positive correlations in southern Europe. The centre of negative correlations over

the North Atlantic in JJA also is displayed for all the systems, but with slightly different values and locations. The seasonal predictions of SLP for SON (Figure 4, fourth column) are mostly negative for the five systems, showing the limitations of the seasonal prediction systems to provide skilful SLP forecasts for that specific season.

These differences in the correlation values across seasons and areas illustrate the difficulties in the selection of a seasonal forecast system that can be used for the development of a tailored seasonal forecast product of SLP-based atmospheric patterns for the wind energy users. There are several seasons and regions where the systems agree in the skilful forecasts of SLP. However, there is no single system that systematically provides better seasonal forecasts of the SLP in the full domain, as illustrated by the maximum correlation in each individual grid point Figure S12. Consequently, as an alternative to convey useful prediction information, the potential of each forecast system to provide forecasts of the frequency of occurrence of the clusters corresponding to the different classifications has been investigated for the four seasons.

The clusters defined in ERA5 have been used to classify the seasonal forecasts of SLP by employing the RMSD assignation method. This method provides classifications (with k ranging from 3 to 6) of the seasonal forecasts with patterns that are, by construction, very similar to those in ERA5 (Figure S13 shows that the spatial correlations between them exceed 0.96 in all the seasons). The frequency of occurrence of the patterns obtained for the seasonal forecasts and the ERA5 patterns have been compared in terms of Pearson correlation in DJF and JJA (Figure 5). Correlations for MAM and SON can be found in the supporting information (Figure S14).

Based on the spatial correlations (Figure 2) and the potential of these clusters to characterize the variability of the wind speed or temperature (Figure 3), $k = 4$ would be the best option to classify the ERA5 SLP daily anomalies for all the seasons. For that specific number of clusters (Figure 5c) in DJF, the DWD-S2 is the only system that provides positive correlations for the four clusters, showing the ability of this system to simulate the year-to-year variations of these clusters 1 month ahead. These positive correlations for DWD-S2 are found for the different number of partitions (although for $k = 5$ and 6 the correlations are not significant for some of the clusters) and it might be related to the high skill shown by this system in DJF (Figure S12). The other four systems also show some of the clusters with positive correlations. For JJA (Figure 5) no system shows positive correlations for all the clusters, but the frequencies of occurrence in the UKMO-GS5 system show positive correlations for all the clusters except the first one. This is related to the higher skill of the UKMO-GS5 SLP in the northern North

Atlantic for that particular season (Figure 4). The correlations in the other systems are mostly negative.

In the equinoctial seasons (Figure S14), the clusters from the MF-S6 system have frequencies of occurrence with a good association with the corresponding frequencies for the clusters in ERA5, as indicated by the positive correlations in most of the clusters for both seasons. In MAM, positive correlations are obtained in most of the clusters for the different systems. However, in SON all the systems but MF-S6 show negative correlations (Figure S14, right column). This is related to the skill of the MF-S6 system in reproducing the SLP in SON (Figure 4), which is higher than for the other systems.

These results show that the ability of the seasonal forecast systems in simulating the year-to-year variations in the frequency of occurrence of the different clusters is more related to the skill of these systems to predict the SLP rather than the number of clusters considered.

4 | SUMMARY AND CONCLUSIONS

The occurrence of specific atmospheric conditions can lead to fluctuations in the wind power production and electricity demand over Europe (Brayshaw *et al.*, 2011; van der Wiel *et al.*, 2019) that can have financial consequences for this sector (Curtis *et al.*, 2016). This connection has recently raised the interest of wind energy users in the atmospheric patterns and their frequency of occurrence. Cluster analysis has been used to describe the atmospheric circulation in winter and summer, but the examples providing classifications for the full-year round are still limited (e.g., Vrac *et al.*, 2014; Grams *et al.*, 2017; Cortesi *et al.*, 2019). Consequently, this paper tried to define a user-oriented classification for each season that can satisfy the needs of the wind energy sector and, at the same time, to illustrate the challenges involved in the development of this aspect of the climate service.

Several methodologies can be employed to classify the atmospheric circulation in a set of a few recurrent patterns. In this study, we have applied the k -means clustering method. The need to define the number of partitions prior to the application of the k -means introduces an important uncertainty. Classifications with three, four, five and six clusters have been explored and compared using a number of criteria for each season. The spatial similarity of the patterns defined for the ERA5 SLP, their ability to explain the variability of the wind speed and temperature, which are essential climate variables for the wind energy activities, and the skill of five state-of-the-art

seasonal forecast systems in predicting their frequency of occurrence are the three criteria considered.

The objective is to identify the minimum number of clusters needed to characterize the atmospheric conditions over the Euro-Atlantic region for each specific season. The clustering approach provides a number of characteristic patterns, but the significance of such patterns is a nontrivial problem (e.g., Stephenson *et al.*, 2004; Christiansen, 2007) and has to be considered with caution. The statistical significance of the different number of clusters considered here has been explored in terms of spatial correlation. The analysis of the correlations between the patterns for the different number of partitions (k) has shown that using four clusters is a good option for all the seasons except for JJA. For summer $k = 3$ can also be a good choice from the point of view of the spatial correlation between clusters. However, in DJF, MAM and SON, five clusters ($k = 5$) could also be considered, as the positive and significant spatial correlations between some clusters are very low.

The choice of a number of clusters has also been considered using a second criterion. We have evaluated the usefulness of each classification to reconstruct the ERA5 wind speed and temperature interannual variability for each grid point over Europe. This reconstruction method is appropriate to explore in which regions the overall influence of the clusters is high and, consequently, more useful to understand the fluctuations of the wind energy resources. Nevertheless, the adequacy of other reconstruction methodologies (Deser *et al.*, 2016; Fereday *et al.*, 2018) to enhance the representation of the year-to-year variations of the wind energy resources by means of atmospheric circulation patterns could be explored in the future. The effectiveness of the reconstruction of wind speed and temperature increases when the number of clusters changes from three to four, which suggests that $k = 4$ is a good selection across the four seasons. However, this choice could be modified depending on the users' needs. For example, it can be seen that a general improvement (i.e., significant increase of the correlation between the reconstructed and the original variable) is obtained in the reconstructed wind speed over southern Europe and Scandinavia when the number of clusters increases from four to five in MAM. Therefore, if the wind energy users request information for a wind farm that is located in that specific location, the number of clusters might be changed. This is not the only example of how this methodology could be adapted to specific applications. Despite the reconstruction being applied to the wind speed and temperature, different indicators such as capacity factor or extreme wind speeds (e.g., 90th percentile) could be also explored, as this

information is often relevant for wind energy decision-making processes (Cannon *et al.*, 2015; Lledó *et al.*, 2019).

Additionally, this work has employed five C3S seasonal forecast systems to investigate the skill of the frequency of occurrence of the atmospheric patterns defined for each k . This assessment is the third criterion considered to select the number of clusters. This information is important as the anticipation of the expected frequency of occurrence of these atmospheric patterns could help to minimize the effects of the atmospheric variability in the wind energy activities. The skill differences between the three, four, five and six clusters considered is not very high, which demonstrates that the skill of a specific classification cannot be attributed to the choice of a specific k . The seasonal prediction systems show a certain degree of skill in some specific clusters depending on the system or the season. For example, it is shown that the DWD-S2 is able to simulate the year-to-year variations of the frequency of occurrence of the different atmospheric patterns in DJF, but in JJA the UKMO-GS5 is the system that provides the highest skill for some specific cluster. This indicates that more than one seasonal forecast system needs to be considered for the development of a user-oriented seasonal forecast product of atmospheric circulation patterns. Nevertheless, the skill of the seasonal predictions in simulating the frequency of occurrence of each cluster one-month ahead is still very limited, with only a few clusters showing positive and significant correlations for specific seasons. This skill is directly associated with the skill found for the SLP, with the systems showing a higher SLP regional skill having higher skill for the cluster frequency. Nevertheless, there are additional factors that can also play a role on the low skill in that area, such as the short hindcast length (1993–2016) or the limited ensemble size (from 40 to 25 members depending on the system) of the C3S seasonal forecast systems available (Scaife *et al.*, 2014; Manzananas *et al.*, 2019; Lledó *et al.*, 2020). The sensitivity of the skill to simulate the frequency of occurrence of the clusters depending on the ensemble size and the hindcast period has been tested (Figure S15) for the ECMWF-S5 seasonal forecast system. This system is suitable for this analysis because it has a long hindcast period (1981–2016) and 51 ensemble members for the start dates of February, May, August and November. The seasonal forecasts of the SLP in DJF (one-month lead) show the highest skill in when 51-ensemble members and 36 winters are considered (Figure S15a). The impact of the hindcast length and ensemble size on the correlation of the frequencies (Figure S11e) shows that the largest hindcast length and ensemble size lead to high correlation values for each k . Furthermore, the maximum

correlations for the different classifications are obtained when 51-ensemble members are used which evidences the need of large ensembles to produce skilful forecasts in the Euro-Atlantic region (e.g., Baker *et al.*, 2018). This analysis also suggests that a $k = 4$ choice might lead to more skilful results in winter, provided that the ensemble size is large enough.

In conclusion, this work has illustrated the main challenges for the development of a set of clusters that can be used for the development of a climate service for the wind industry. Despite the fact that previous literature has already shown that four clusters are the optimal choice for the description of the Euro-Atlantic climate variability in winter and summer, this work also explores the optimal number of clusters in the transition seasons. We have shown that four clusters are a good option to describe the atmospheric variability in the Euro-Atlantic domain, particularly when these patterns are used to understand the year-to-year fluctuations of the surface wind speed and temperature. The analysis of the different options in the classification has revealed that there is not an optimal number of clusters that can lead to generally skilful forecasts of the frequency of occurrence of the atmospheric patterns. The skill of the predicted frequencies of occurrence depends much more on the underlying sea level pressure skill, which depends on the seasonal forecast system considered. Therefore, the use of more than one seasonal forecast system is recommended for the development of this type of climate services product.

Future research will focus on the application of the strategy followed in this work for the identification of clusters that can be useful for the wind energy sector in different regions, like North America, which is a very important area for the wind energy business. Furthermore, the comparison of the forecast results with those that could be obtained with alternative classifications, such as the yearly classification proposed by (Grams *et al.*, 2017), the modified version proposed by (Garrido-Perez *et al.*, 2020) or the targeted circulation types defined by (Bloomfield *et al.*, 2020), can also be interesting from the wind energy point of view. Finally, the usefulness of the atmospheric circulation patterns to understand the variability of tailored indicators such as extreme wind speeds or wind power generation will be investigated.

ACKNOWLEDGEMENTS

This research has been funded by the S2S4E (GA 776787) Horizon 2020 project, the Ministerio de Ciencia, Innovación y Universidades as part of the CLINSA project (CGL2017-85791-R) and the Juan de la Cierva – Incorporación Grant (IJCI-2016-29776). The analyses and plots of this work have been performed with the

s2dverification (Manubens *et al.*, 2018), CStools (<https://CRAN.R-project.org/package=CStools>) and startR (<https://CRAN.R-project.org/package=startR>) R-language-based software packages. Finally, we would like to thank Pierre-Antoine Bretonnière, Margarida Samsó, Nicolau Manubens and Núria Pérez-Zanón for their technical support at different stages of this project. We also acknowledge the two anonymous reviewers for their useful comments.

ORCID

Verónica Torralba  <https://orcid.org/0000-0002-8941-1548>

REFERENCES

- Baker, L., Shaffrey, L., Sutton, R., Weisheimer, A. and Scaife, A. (2018) An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic oscillation in multimodel seasonal forecasts. *Geophysical Research Letters*, 45(15), 7808–7817. <https://doi.org/10.1029/2018GL078838>.
- Beck, C., Philipp, A. and Streicher, F. (2016) The effect of domain size on the relationship between circulation type classifications and surface climate. *International Journal of Climatology*, 36(7), 2692–2709. <https://doi.org/10.1002/joc.3688>.
- Bessec, M. and Fouquau, J. (2008) The non-linear link between electricity consumption and temperature in europe: a threshold panel approach. *Energy Economics*, 30(5), 2705–2721. <https://doi.org/10.1016/j.eneco.2008.02.003>.
- Bloomfield, H.C., Brayshaw, D.J. and Charlton-Perez, A.J. (2020) Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types. *Meteorological Applications*, 27(1), e1858. <https://doi.org/10.1002/met.1858>.
- Brayshaw, D.J., Troccoli, A., Fordham, R. and Methven, J. (2011) The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: a case study over the UK. *Renewable Energy*, 36(8), 2087–2096. <https://doi.org/10.1016/j.renene.2014.10.024>.
- C3S, 2017. *ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate*. <https://cds.climate.copernicus.eu/cdsapp#!/home>.
- Cannon, D., Brayshaw, D., Methven, J., Coker, P. and Lenaghan, D. (2015) Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. *Renewable Energy*, 75, 767–778. <https://doi.org/10.1016/j.renene.2014.10.024>.
- Cassou, C. (2008) Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature*, 455(7212), 523–527. <https://doi.org/10.1038/nature07286>.
- Cassou, C., Terray, L. and Phillips, A.S. (2005) Tropical Atlantic influence on European heat waves. *Journal of Climate*, 18(15), 2805–2811. <https://doi.org/10.1175/JCLI3506.1>.
- Ceglar, A., Turco, M., Toreti, A. and Doblas-Reyes, F.J. (2017) Linking crop yield anomalies to large-scale atmospheric circulation in Europe. *Agricultural and Forest Meteorology*, 240, 35–45. <https://doi.org/10.1016/j.agrformet.2017.03.019>.

- Christiansen, B. (2007) Atmospheric circulation regimes: can cluster analysis provide the number? *Journal of Climate*, 20(10), 2229–2250. <https://doi.org/10.1175/JCLI4107.1>.
- Cleveland, W.S. and Devlin, S.J. (1988) Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596–610. <https://doi.org/10.1080/01621459.1988.10478639>.
- Cortesi, N., Torralba, V., González-Reviriego, N., Soret, A. and Doblas-Reyes, F.J. (2019) Characterization of European wind speed variability using weather regimes. *Climate Dynamics*, 53, 1–16. <https://doi.org/10.1007/s00382-019-04839-5>.
- Curtis, J., Lynch, M.Á. and Zubiate, L. (2016) The impact of the North Atlantic oscillation on electricity markets: a case study on Ireland. *Energy Economics*, 58, 186–198. <https://doi.org/10.1016/j.eneco.2016.07.003>.
- Dawson, A. and Palmer, T. (2015) Simulating weather regimes: impact of model resolution and stochastic parameterization. *Climate Dynamics*, 44(7–8), 2177–2193. <https://doi.org/10.1007/s00382-014-2238-x>.
- Dawson, A., Palmer, T.N. and Corti, S. (2012) Simulating regime structures in weather and climate prediction models. *Geophysical Research Letters*, 39(21), L21805. <https://doi.org/10.1029/2012GL053284>.
- De Cian, E., Lanzi, E. and Roson, R. (2013) Seasonal temperature variations and energy demand. *Climatic Change*, 116(3–4), 805–825. <https://doi.org/10.1007/s10584-012-0514-5>.
- De Felice, M., Alessandri, A. and Catalano, F. (2015) Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy*, 137, 435–444. <https://doi.org/10.1016/j.apenergy.2014.10.030>.
- Deser, C., Terray, L. and Phillips, A.S. (2016) Forced and internal components of winter air temperature trends over north america during the past 50 years: mechanisms and implications. *Journal of Climate*, 29(6), 2237–2258. <https://doi.org/10.1175/JCLI-D-15-0304.1>.
- Dorel, L., Ardilouze, C., Déqué, M., Batté, L. and Guérémy, J. F. 2017 *Documentation of the Météo-France Pre-Operational Seasonal Forecasting System*. <http://www.umr-cnrm.fr/IMG/pdf/system6-technical.pdf>.
- DWD, 2020. *German Climate Forecast System*. https://www.dwd.de/EN/ourservices/seasonals_forecasts/project_description.html?nn=641552&lsbId=61978.
- Fereday, D., Chadwick, R., Knight, J. and Scaife, A.A. (2018) Atmospheric dynamics is the largest source of uncertainty in future winter European rainfall. *Journal of Climate*, 31(3), 963–977. <https://doi.org/10.1175/JCLI-D-17-0048.1>.
- Fereday, D.R., Knight, J.R., Scaife, A.A., Folland, C.K. and Philipp, A. (2008) Cluster analysis of North Atlantic–European circulation types and links with tropical Pacific Sea surface temperatures. *Journal of Climate*, 21(15), 3687–3703. <https://doi.org/10.1175/2007JCLI1875.1>.
- Ferranti, L. and Corti, S. (2011) New clustering products. *ECMWF Newsletter*, 127(6–11), 1–2. <https://doi.org/10.21957/lr3bcise>.
- Ferranti, L., Corti, S. and Janousek, M. (2015) Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, 141(688), 916–924. <https://doi.org/10.1002/qj.2411>.
- García-Morales, M.B. and Dubus, L. (2007) Forecasting precipitation for hydroelectric power management: how to exploit GCM's seasonal ensemble forecasts. *International Journal of Climatology*, 27(12), 1691–1705. <https://doi.org/10.1002/joc.1608>.
- Garrido-Perez, J.M., Ordóñez, C., Barriopedro, D., García-Herrera, R. and Paredes, D. (2020) Impact of weather regimes on wind power variability in western Europe. *Applied Energy*, 264(114), 731. <https://doi.org/10.1016/j.apenergy.2020.114731>.
- Grams, C.M., Beerli, R., Pfenninger, S., Staffell, I. and Wernli, H. (2017) Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, 7(8), 557–562. <https://doi.org/10.1038/nclimate3338>.
- Gueye, A., Janicot, S., Niang, A., Sawadogo, S., Sultan, B., Diongue-Niang, A. and Thiria, S. (2012) Weather regimes over Senegal during the summer monsoon season using self-organizing maps and hierarchical ascendant classification. Part I: synoptic time scale. *Climate Dynamics*, 36(1–2), 1–18. <https://doi.org/10.1007/s00382-010-0782-6>.
- Hall, R.J. and Hanna, E. (2018) North Atlantic circulation indices: links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. *International Journal of Climatology*, 38, e660–e677. <https://doi.org/10.1002/joc.5398>.
- Hannachi, A., Straus, D.M., Franzke, C.L.E., Corti, S. and Woollings, T. (2017) Low frequency nonlinearity and regime behavior in the northern hemisphere extra-tropical atmosphere. *Reviews of Geophysics*, 55, 199–234. <https://doi.org/10.1002/2015RG000509>.
- Hartigan, J.A. and Wong, M.A. (1979) Algorithm AS 136: a k-means clustering algorithm. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 28(1), 100–108. <https://doi.org/10.2307/2346830>.
- Hartmann, D.L., et al. (2013) Observations: atmosphere and surface. In: *Climate change 2013 the physical science basis: working group I contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, United Kingdom and New York, NY USA: Cambridge University Press, pp. 159–254.
- Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J. and Tveito, O.E. (2008) Classifications of atmospheric circulation patterns: recent advances and applications. *Annals of the New York Academy of Sciences*, 1146(1), 105–152. <https://doi.org/10.1196/annals.1446.019>.
- Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremier, D., Weisheimer, A., Balsamo, G., Keeley, S.P.E., Mogensen, K., Zuo, H. and Monge-Sanz, B.M. (2019) SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3), 1087–1117. <https://doi.org/10.5194/gmd-12-1087-2019>.
- Jolliffe, I.T. and Philipp, A. (2010) Some recent developments in cluster analysis. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9–12), 309–315. <https://doi.org/10.1016/j.pce.2009.07.014>.
- Jolliffe, I.T. and Stephenson, D.B. (2012) *Forecast Verification: A practitioner's Guide in Atmospheric Science*. Chichester, United Kingdom: John Wiley & Sons.
- Lavaysse, C., Vogt, J., Toreti, A., Carrera, M.L. and Pappenberger, F. (2018) On the use of weather regimes to forecast meteorological drought over Europe. *Natural Hazards and*

- Earth System Sciences*, 18(12), 3297–3309. <https://doi.org/10.5194/nhess-18-3297-2018>.
- Lledó, L., Cionni, I., Torralba, V., Bretonnière, P.-A. and Samsó, M. (2020) Seasonal prediction of Euro-Atlantic teleconnections from multiple systems. *Environmental Research Letters*, 15(7), 074009. <https://doi.org/10.1088/1748-9326/ab87d2/meta>.
- Lledó, L., Torralba, V., Soret, A., Ramon, J. and Doblas-Reyes, F.J. (2019) Seasonal forecasts of wind power generation. *Renewable Energy*, 143, 91–100. <https://doi.org/10.1016/j.renene.2019.04.135>.
- Lowe, R., Garca-Dez, M., Ballester, J., Creswick, J., Robine, J.-M., Herrmann, F.R. and Rodó, X. (2016) Evaluation of an early-warning system for heat wave-related mortality in Europe: implications for sub-seasonal to seasonal forecasting and climate services. *International Journal of Environmental Research and Public Health*, 13(2), 206. <https://doi.org/10.3390/ijerph13020206>.
- Mahlstein, I., Spirig, C., Liniger, M.A. and Appenzeller, C. (2015) Estimating daily climatologies for climate indices derived from climate model data and observations. *Journal of Geophysical Research: Atmospheres*, 120(7), 2808–2818. <https://doi.org/10.1002/2014JD022327>.
- Manubens, N., Caron, L.P., Hunter, A., Bellprat, O., Exarchou, E., Fučkar, N.S., Garcia-Serrano, J., Massonnet, F., Ménégos, M., Sicardi, V., Batté, L., Prodhomme, C., Torralba, V., Cortesi, N., Mula-Valls, O., Serradell, K., Guemas, V. and Doblas-Reyes, F. J. (2018) An R package for climate forecast verification. *Environmental Modelling & Software*, 103, 29–42. <https://doi.org/10.1016/j.envsoft.2018.01.018>.
- Manzanas, R., Gutiérrez, J.M., Bhend, J., Hemri, S., Doblas-Reyes, F.J., Torralba, V., Penabad, E. and Brookshaw, A. (2019) Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset. *Climate Dynamics*, 53, 1287–1305. <https://doi.org/10.1007/s00382-019-04640-4>.
- Michelangeli, P.-A., Vautard, R. and Legras, B. (1995) Weather regimes: recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, 52(8), 1237–1256. [https://doi.org/10.1175/1520-0469\(1995\)052<1237:WRRASQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1237:WRRASQ>2.0.CO;2).
- Neal, R., Fereday, D., Crocker, R. and Comer, R.E. (2016) A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorological Applications*, 23(3), 389–400. <https://doi.org/10.1002/met.1563>.
- Palin, E.J., Scaife, A.A., Wallace, E., Pope, E.C., Arribas, A. and Brookshaw, A. (2016) Skillful seasonal forecasts of winter disruption to the UK transport system. *Journal of Applied Meteorology and Climatology*, 55(2), 325–344. <https://doi.org/10.1175/JAMC-D-15-0102.1>.
- Palmer, T.N. and Anderson, D.L. (1994) The prospects for seasonal forecasting—a review paper. *Quarterly Journal of the Royal Meteorological Society*, 120(518), 755–793. <https://doi.org/10.1002/qj.49712051802>.
- Quesada, B., Vautard, R., Yiou, P., Hirschi, M. and Seneviratne, S.I. (2012) Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change*, 2(10), 736–741. <https://doi.org/10.1038/nclimate1536>.
- Reyers, M., Pinto, J.G. and Moemken, J. (2015) Statistical-dynamical downscaling for wind energy potentials: evaluation and applications to decadal hindcasts and climate change projections. *International Journal of Climatology*, 35(2), 229–244. <https://doi.org/10.1002/joc.3975>.
- Sanna, A., Borrelli, A., Athanasiadis, P., Materia, S., Storto, A., Navarra, A., Tibaldi, S. and Gualdi, S., 2017: *CMCC-SPS3: The CMCC Seasonal Prediction System 3*. <https://www.cmcc.it/wp-content/uploads/2017/12/rp0285-csp-10-2017.pdf>.
- Scaife, A., et al. (2014) Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, 41 (7), 2514–2519. <https://doi.org/10.1002/2014GL059637>.
- Soret, A., et al. (2019) Sub-seasonal to seasonal climate predictions for wind energy forecasting. *Journal of Physics: Conference Series*, 1222, 012009. <https://doi.org/10.1088/1742-6596/1222/1/012009>.
- Staffell, I. and Pfenninger, S. (2018) The increasing impact of weather on electricity supply and demand. *Energy*, 145, 65–78. <https://doi.org/10.1016/j.energy.2017.12.051>.
- Stephenson, D.B., Hannachi, A. and O'Neill, A. (2004) On the existence of multiple climate regimes. *Quarterly Journal of the Royal Meteorological Society*, 130(597), 583–605. <https://doi.org/10.1256/qj.02.146>.
- Straus, D.M., Corti, S. and Molteni, F. (2007) Circulation regimes: chaotic variability versus SST-forced predictability. *Journal of Climate*, 20(10), 2251–2272. <https://doi.org/10.1175/JCLI4070.1>.
- Stryhal, J. and Huth, R. (2017) Classifications of winter Euro-Atlantic circulation patterns: an intercomparison of five atmospheric reanalyses. *Journal of Climate*, 30(19), 7847–7861. <https://doi.org/10.1175/JCLI-D-17-0059.1>.
- Thornton, H.E., Scaife, A.A., Hoskins, B.J. and Brayshaw, D.J. (2017) The relationship between wind power, electricity demand and winter weather patterns in Great Britain. *Environmental Research Letters*, 12(6), 064 017. <https://doi.org/10.1088/1748-9326/aa57ab>.
- Torralba, V., 2019 Seasonal climate prediction for the wind energy sector: methods and tools for the development of a climate service. Ph.D. thesis, Universidad Complutense de Madrid. <https://eprints.ucm.es/56841/>.
- Torralba, V., Doblas-Reyes, F.J., MacLeod, D., Christel, I. and Davis, M. (2017) Seasonal climate prediction: a new source of information for the management of wind energy resources. *Journal of Applied Meteorology and Climatology*, 56(5), 1231–1247. <https://doi.org/10.1175/JAMC-D-16-0204.1>.
- van der Wiel, K., Bloomfield, H.C., Lee, R.W., Stoop, L.P., Blackport, R., Screen, J.A. and Selten, F.M. (2019) The influence of weather regimes on european renewable energy production and demand. *Environmental Research Letters*, 14(9), 094 010. <https://doi.org/10.1088/1748-9326/ab38d3>.
- Vautard, R. (1990) Multiple weather regimes over the North Atlantic: analysis of precursors and successors. *Monthly Weather Review*, 118(10), 2056–2081. [https://doi.org/10.1175/1520-0493\(1990\)118<2056:MWOTN>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<2056:MWOTN>2.0.CO;2).
- Vrac, M., Vaittinada Ayar, P. and Yiou, P. (2014) Trends and variability of seasonal weather regimes. *International Journal of Climatology*, 34(2), 472–480. <https://doi.org/10.1002/joc.3700>.
- Walz, M.A., Befort, D.J., Kirchner-Bossi, N.O., Ulbrich, U. and Leckebusch, G.C. (2018) Modelling serial clustering and inter-annual variability of European winter windstorms based on

- large-scale drivers. *International Journal of Climatology*, 38(7), 3044–3057. <https://doi.org/10.1002/joc.5481>.
- Williams, K., et al. (2015) The met office global coupled model 2.0 (gc2) configuration. *Geoscientific Model Development*, 8(5), 1509–1524. <https://doi.org/10.5194/gmd-8-1509-2015>.
- Yiou, P., Goubanova, K., Li, Z.X. and Nogaj, M. (2008) Weather regime dependence of extreme value statistics for summer temperature and precipitation. *Nonlinear Processes in Geophysics*, 15(3), 365–378. <https://doi.org/10.5194/npg-15-365-2008>.
- Zubiate, L., McDermott, F., Sweeney, C. and O'Malley, M. (2017) Spatial variability in winter NAO–wind speed relationships in western Europe linked to concomitant states of the East Atlantic and Scandinavian patterns. *Quarterly Journal of the Royal Meteorological Society*, 143(702), 552–562. <https://doi.org/10.1002/qj.2943>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Torralba V, Gonzalez-Reviriego N, Cortesi N, *et al.* Challenges in the selection of atmospheric circulation patterns for the wind energy sector. *Int J Climatol*. 2020;1–17. <https://doi.org/10.1002/joc.6881>